

Analysing surveys of our Galaxy – II. Determining the potential

Paul J. McMillan^{*} and James J. Binney

Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford, OX1 3NP, UK

24 May 2013

ABSTRACT

We consider the problem of determining the Galaxy’s gravitational potential from a star catalogue. We show that orbit-based approaches to this problem suffer from unacceptable numerical noise deriving from the use of only a finite number of orbits. An alternative approach, which requires an ability to determine the model’s phase-space density at predetermined positions and velocities, has a level of numerical noise that lies well below the intrinsic uncertainty associated with the finite size of the catalogue analysed. A catalogue of 10 000 stars brighter than $V = 17$ and distributed over the sky at $b > 30$ degrees enables us to determine the scaleheight of the disc that contributes to the potential with an uncertainty below 20 pc if the catalogue gives proper motions, line-of-sight velocities and parallaxes with errors typical of the Gaia Catalogue, rising to 36 pc if only proper motions are available. The uncertainty in the disc’s scalelength is significantly smaller than 0.25 kpc.

Key words: Galaxy: kinematics and dynamics – Galaxy: structure – methods: data analysis

1 INTRODUCTION

Very large resources are currently being devoted to surveys of our Galaxy, both from the ground (e.g. APOGEE, RAVE and Gaia-ESO: Eisenstein et al. 2011; Steinmetz et al. 2006; Gilmore et al. 2012) and from space (e.g. Gaia: Perryman et al. 2001). These surveys are being undertaken in the expectation that they will reveal the current kinematic and chemical structure of our Galaxy and how our Galaxy was assembled. The latter is a pivotal question for cosmology since our Galaxy is typical of the galaxies that dominate star formation in the contemporary Universe, and the consensus Λ CDM cosmology has given us a moderately clear view of how such galaxies formed.

Since stars and dark matter orbit freely in the Galaxy’s gravitational potential, our understanding of the Galaxy can never be better than our knowledge of its gravitational potential $\Phi(\mathbf{x})$. Since even at the Sun we are unable to measure the density of dark matter, Φ cannot be determined from Poisson’s equation, but must be constrained by measuring the motions of objects that we can see – in practice stars and interstellar gas. Historically, the crucial data have been the circular speed in the plane, $v_c(R)$, estimated from the line-of-sight velocities v_{\parallel} of interstellar gas (Malhotra 1994a), the local density $\rho(R_0)$ estimated from the kinematics of nearby stars (Creze et al. 1998; Holmberg & Flynn

2004), the local surface density $\Sigma_{1.1}$ estimated from observations of stars a few hundred parsecs away from the plane (Kuijken & Gilmore 1989), and the proper motion of Sagittarius A* at the Galactic centre (Reid & Brunthaler 2004; Gillessen et al. 2009). One can find models of the Galactic potential which fit all of these constraints (e.g. Dehnen & Binney 1998; Klypin, Zhao, & Somerville 2002; McMillan 2011), but unfortunately they do not constrain the density of dark matter very strongly. The data that are now available to us, or will shortly become available, open new horizons in the determination of Φ .

The questions we address are:

- given a star catalogue, what is the best way to constrain Φ ?
- When we proceed in the optimal way, how strongly do data of the type that will shortly be available constrain Φ ?

The data available from surveys of the Galaxy fundamentally differs from that generally available for external galaxies. It is practical to determine not just the position on the sky and line-of-sight velocity of a star, but also proper motions and distance from the Sun with a useful degree of accuracy. However, there is a strong bias in which stars are observed by the survey – those sufficiently close to the Sun’s position in the Galaxy and which lie within the selection function of the survey. For observations of external galaxies there is no such bias created by the Sun’s position, and full surface brightness profiles can be found, but the velocity in-

^{*} E-mail: p.mcmillan1@physics.ox.ac.uk

formation is generally limited to binned line-of-sight velocity distributions. Ensuring that models can accurately represent data of the precision available for stars in the Galaxy is a significant problem.

A compelling argument can be made that the availability of models of sufficient sophistication is the key to extracting science from a star catalogue (McMillan & Binney 2012, henceforth Paper I). One must discriminate between various model Galaxies by inferring their a posteriori probabilities, given the catalogue. The simplest possible dynamical model Galaxy is defined by the pair of functions (f, Φ) , where $f(\mathbf{x}, \mathbf{v})$ is the distribution function (DF).

When estimating Φ an unavoidable assumption is that the Galaxy is in a steady state: without this assumption the observed kinematics of any tracer objects are consistent with any potential; it is only by assuming that Φ is deep enough to prevent a tracer population expanding, and not so deep as to cause contraction in the next dynamical time that we can constrain Φ . The assumption that our tracers are in dynamical equilibrium permits us to invoke the strong Jeans theorem and conclude that the DF f of our tracer population is a function $f(I_1, I_2, \dots)$ of whatever isolating integrals Φ admits. In the 1970s numerical experiments revealed that typical galaxy potentials admit three independent isolating integrals, for example, energy $E = \frac{1}{2}v^2 + \Phi$, the component of angular momentum L_z about the potential's symmetry axis and a "third integral" I_3 , which controls the division of energy in excess of the energy $E_c(L_z)$ of a circular orbit of angular momentum L_z between oscillation in radius and oscillation perpendicular to the potential's equatorial plane.

Any function of isolating integrals is itself an isolating integral, so there is considerable flexibility in the choice of arguments of the DF. We have argued elsewhere (McMillan & Binney 2008; Binney 2010) that there are compelling reasons to choose the actions J_r , $J_\phi \equiv L_z$ and J_z as our isolating integrals. So we work with these integrals here and take a model galaxy to be the pair of functions $(f(\mathbf{J}), \Phi(\mathbf{x}))$. However, working with alternative integrals would not change our conclusions in any essential way; alternative integrals would simply make the computations harder and less transparent. In particular, our arguments apply to models of the type that are most widely used in studies of external galaxies: Schwarzschild (1979) models. These models comprise a potential Φ and an orbit library, each element of which is the time series of phase-space coordinates $(\mathbf{x}(t), \mathbf{v}(t))$ obtained by integrating the equations of motion in $\Phi(\mathbf{x})$ for a particular initial condition, and a weight $w \geq 0$ with which that time series is employed in the model. In effect these models use the initial conditions of integrations as isolating integrals, and the weights w are surrogates for the value of the DF on the given initial conditions.

In general terms the procedure for modelling a catalogue is to determine the probability of the data given some pair (f, Φ) and then to use Bayes' theorem to convert the probability of the data into the probability of the pair. Ideally, the probabilities of every plausible pair (f, Φ) would be determined, but in practice one has to be content with a search for a limited number of more likely pairs. The problem is made computationally tractable by considering each candidate potential in turn, and then finding the most probable companion DF. In Paper I we showed that when a catalogue of 10 000 stars is constructed in a known potential from a DF

of a given functional form, the DF can be recovered to good precision from the catalogue. In this paper we explore our ability to determine the correct potential by repeating the DF-fitting step for a series of potentials and identifying the potential that yields the largest likelihood for the catalogue.

We show that this problem cannot be efficiently solved by any orbit-based technique such as N-body modelling, Schwarzschild modelling, or torus modelling of the type used in Paper I. We show further that the problem can be solved if we have available expressions for the isolating integrals as functions of the conventional phase-space variables, for example $\mathbf{J}(\mathbf{x}, \mathbf{v})$. This second approach was adopted by Ting et al. (2012), and we extend it to include both observational errors and realistic selection effects, and to exploit the more powerful approach to the determination of actions of Binney (2012a).

The outline of the paper is as follows. In Section 2 we describe the catalogues that we consider, the models that we compare them to, and the tools used to do the comparison. In Section 3 we explain the two competing methods we use to analyse the data. In Section 4 we show that orbit-based methods are ill-suited to this analysis. In Section 5 we demonstrate methods using expressions for $\mathbf{J}(\mathbf{x}, \mathbf{v})$ that are capable of successfully determining the true potential.

2 THEORETICAL FRAMEWORK

Three actions J_i and three conjugate angle coordinates θ_i provide exceptionally convenient coordinates for objects orbiting in a stationary gravitational potential Φ . The actions are conserved quantities and the angles increase linearly with time, $\theta_i(t) = \theta_i(0) + \Omega_i(\mathbf{J})t$, where Ω_i is a frequency.

Thus \mathbf{J} labels an orbit and $\boldsymbol{\theta}$ specifies a point on that orbit. The usual phase space coordinates \mathbf{x}, \mathbf{v} are 2π -periodic in each angle coordinate θ_i . Angle-action coordinates have the convenient property that

$$\left| \frac{\partial(\boldsymbol{\theta}, \mathbf{J})}{\partial(\mathbf{x}, \mathbf{v})} \right| = 1, \quad (1)$$

so it is simple to relate a density in angle-action space to the phase-space density $f(\mathbf{x}, \mathbf{v})$.

The relationship $(\boldsymbol{\theta}, \mathbf{J}) \leftrightarrow (\mathbf{x}, \mathbf{v})$ depends upon the gravitational potential Φ . Unfortunately analytical expressions for this relationship are only known for a very limited set of potentials. In recent years a great deal of effort has gone into developing numerical approximations to this relationship (McMillan & Binney 2008; Binney 2010; Binney & McMillan 2011; Sanders 2012a; Binney 2012a). In this paper we first use the machinery described in McMillan & Binney (2008) that yields $(\mathbf{x}(\boldsymbol{\theta}, \mathbf{J}), \mathbf{v}(\boldsymbol{\theta}, \mathbf{J}))$ and then the machinery described in Binney (2012a), which gives the inverse transformation $(\boldsymbol{\theta}(\mathbf{x}, \mathbf{v}), \mathbf{J}(\mathbf{x}, \mathbf{v}))$.

2.1 Torus modelling

Torus modelling (McMillan & Binney 2008, and references therein) is a method which, for a single value \mathbf{J} in a given potential Φ , provides (through a numerical minimisation) an expression for the phase-space coordinates (\mathbf{x}, \mathbf{v}) in terms of $\boldsymbol{\theta}$. Thus it tells us the complete phase-space structure of the specified orbit \mathbf{J} . We refer to this model of an orbit as a

“torus” because the three-dimensional surface mapped out in phase space as the θ_i vary over their full range $(0, 2\pi)$ is isomorphic to a 3-torus.

Thus with this method it is easy to find (\mathbf{x}, \mathbf{v}) in terms of $(\boldsymbol{\theta}, \mathbf{J})$, but far harder to find $(\boldsymbol{\theta}, \mathbf{J})$ given (\mathbf{x}, \mathbf{v}) (though it can be done iteratively, e.g. McMillan & Binney 2008).

Torus modelling is best understood as an extension of Schwarzschild modelling (Schwarzschild 1979) in which time series are replaced by orbital tori. This replacement brings a number of advantages (Binney & McMillan 2011). Torus modelling was the method used to in Paper I, and has also been used in modelling the Hyades moving group (McMillan 2013) and to show how one can disentangle the history of a disrupted satellite object that we observe as debris in the Solar neighbourhood (McMillan & Binney 2008).

2.2 Stäckel approximation

Binney (2012a) introduced an algorithm for calculating the actions of stars with known phase-space coordinates in axisymmetric potentials. It is based upon the approximation that in the region probed by a given orbit, the potential of interest does not differ greatly from a Stäckel potential (e.g. de Zeeuw 1985; Binney & Tremaine 2008, §3.5.3). With this assumption it is possible to estimate the radial and vertical actions J_r and J_z at any point (\mathbf{x}, \mathbf{v}) from one-dimensional integrals over coordinates that form a system of confocal ellipsoidal coordinates. Thus when this apparatus is used, it is easy to obtain $(\boldsymbol{\theta}, \mathbf{J})$ from (\mathbf{x}, \mathbf{v}) but hard to proceed in the opposite direction.

This approximation gives values of \mathbf{J} for typical orbits in the thin disc which are around a factor of 4 more accurate than those found by the “adiabatic approximation” which has been used for the same purpose (e.g. Binney 2010; Ting et al. 2012). It provides an even greater improvement in accuracy for the orbits of many thick-disc stars, to which the adiabatic approximation does not apply. It has been used to fit DFs to observational data for the Solar neighbourhood, given assumed gravitational potentials (Binney 2012b).

2.3 Distribution functions

The distribution functions used in this paper are all based upon the “quasi-isothermal” DF (Binney & McMillan 2011). Here we modify the notation used previously in two respects: (i) we change the normalisation of f so when integrated over all phase space it produces unity, and (ii) we replace the parameter q by $R_\sigma \equiv R_d/q$, where R_σ is the radial scale on which the velocity dispersions decline with increasing radius, and R_d is the conventional scalelength of the approximately isothermal disc.

$$f(\mathbf{J}) \equiv \frac{\Omega\nu\Sigma}{2\pi^2 M \sigma_r^2 \sigma_z^2 \kappa} \bigg|_{R_c} \text{cut}(L_z) e^{-\kappa J_r / \sigma_r^2} e^{-\nu J_z / \sigma_z^2}, \quad (2)$$

with $\mathbf{J} \equiv (J_r, J_z, L_z)$. Here $\Omega(L_z)$ is the circular frequency for angular momentum L_z , $\kappa(L_z)$ is the radial epicycle frequency and $\nu(L_z)$ is its vertical counterpart. $\Sigma(L_z) = \Sigma_0 e^{-R_c/R_d}$ is the (approximate) radial surface-density profile, where $R_c(L_z)$ is the radius of the circular orbit with

angular momentum L_z , and $M = 2\pi\Sigma_0 R_d^2$ is a constant included to ensure that

$$\int d^3\mathbf{J} f(\mathbf{J}) = 1. \quad (3)$$

The factor $\text{cut}(L_z)$ is included to ensure that we do not have equal numbers of stars rotating in each direction. We use

$$\text{cut}(L_z) = \frac{1}{2} [1 + \tanh(L_z/L_0)] \quad (4)$$

where the value of L_0 is unimportant in this study provided it is small compared to the angular momentum of the Sun. We hold it fixed at $L_0 = 10 \text{ kpc km s}^{-1}$. The functions $\sigma_z(L_z)$ and $\sigma_r(L_z)$ control the vertical and radial velocity dispersions. We adopt

$$\begin{aligned} \sigma_r(L_z) &= \sigma_{r0} e^{(R_0 - R_c)/R_\sigma} \\ \sigma_z(L_z) &= \sigma_{z0} e^{(R_0 - R_c)/R_\sigma}, \end{aligned} \quad (5)$$

where σ_{r0} and σ_{z0} are parameters that are set to values close to the radial and vertical velocity dispersions at the Sun. The observed insensitivity to radius of the scaleheights of extragalactic discs suggests $R_\sigma \sim 2R_d$, if R_d is the scalelength of the disc that dominates the potential. To simplify calculations we hold $R_\sigma = R_d/0.45$ (as in Paper I).

Binney (2012b) showed that by superposing a large number of quasi-isothermal DFs, one can obtain a model that is consistent with the local stellar density and velocity distribution as revealed by the Geneva-Copenhagen survey (Holmberg et al. 2009). In this study we, as in Paper I, restrict ourselves to simple two-disc models in order to provide some straightforward demonstrations of the principles involved. These are of the form

$$f(\mathbf{J}) = (1 - \lambda) f_{\text{thin}}(\mathbf{J}) + \lambda f_{\text{thick}}(\mathbf{J}) \quad (6)$$

with f_{thin} and f_{thick} of the form given in equation 2, and λ the fraction in the thick disc. Extending this work to more complicated DFs is, in principle, straightforward.

2.4 Numerical details of models used in this study

The models we use to test our analysis techniques are discrete realisations obtained by sampling a DF that has two quasi-isothermal discs with parameters as listed in Table 2. We refer to this DF as f_{true} . This is identical to one of the DFs used in Paper I and it is sampled using the torus machinery as described in Paper I.

In all cases the model is constructed in the “convenient” Milky Way potential given by McMillan (2011), which we will refer to as the “Mc11” potential. This is an axisymmetric model, in which the potential is assumed to be produced by a Galactic bulge, thin and thick exponential discs, and a spherical Navarro, Frenk, & White (1996) halo. The density of the bulge is

$$\rho_b = \frac{\rho_{b,0}}{(1 + r'/r_0)^\alpha} \exp \left[- (r'/r_{\text{cut}})^2 \right], \quad (7)$$

where, in cylindrical coordinates,

$$r' = \sqrt{R^2 + (z/q)^2} \quad (8)$$

with $\alpha = 1.8$, $r_0 = 0.075 \text{ kpc}$, $r_{\text{cut}} = 2.1 \text{ kpc}$, and axis ratio $q = 0.5$; the densities of the two discs are of the form

$$\rho_d(R, z) = \frac{\Sigma_{d,0}}{2z_d} \exp \left(-\frac{|z|}{z_d} - \frac{R}{R_d} \right), \quad (9)$$

Table 2. Parameters of the DF used to construct our catalogues, f_{true} . The fraction in the thick disc component is $\lambda = 0.23$.

Disc	R_d (kpc)	σ_{r0} (km s ⁻¹)	σ_{z0} (km s ⁻¹)
Thin	3.0	27	20
Thick	3.5	48	44

with scaleheight z_d , scalelength R_d and central surface density $\Sigma_{d,0}$; and the density of the halo is of the form

$$\rho_h = \frac{\rho_{h,0}}{x(1+x)^2}, \quad (10)$$

where $x = r/r_h$, with r_h the scale-radius. The parameters of the Mc11 model are shown in Table 1.

We also consider a range of other potentials, to show how well our analysis techniques work when asked to compare the likelihoods of competing plausible potentials. We perform a number of tests in which all parameters are held constant at those of Mc11, except the disc scaleheights – in these cases the ratio of the two disc scaleheights is held constant, and the quoted value is that of the thin-disc scaleheight.

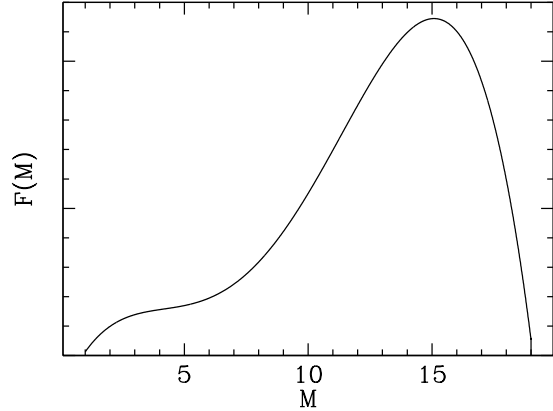
We also perform tests in which the disc scalelengths are varied. Again we hold the ratio of the two disc scalelengths constant. However, now leaving the other parameters of the potential unchanged would lead to substantial changes in the properties of the model (for example in the circular speed), so we constrain the parameters of each potential in the same way as in McMillan (2011), with the Sun’s position and the disc scaleheights being held constant.

The constraints on the potentials employed include the proper motion of Sgr A* (Reid & Brunthaler 2004). Given the Sun’s Galactocentric distance (8.5 kpc) and peculiar motion with respect to the local standard of rest, (Schönrich, Binney, & Dehnen 2010) the proper motion of Sgr A* strongly constrains the local circular speed. We also fit the potentials to the terminal velocity of the ISM at $30^\circ < l < 90^\circ$ (Malhotra 1994b, 1995) and to observed maser sources (e.g. McMillan & Binney 2010) which constrain the shape of the circular-speed curve, and to the vertical force 1.1 kpc from the plane at the Solar radius (Kuijken & Gilmore 1991). For full details of the constraints applied, see McMillan (2011).

Constructing each potential in this way ensures that we are comparing potentials that are all good fits to existing basic kinematic data, which are currently state-of-the-art constraints on the Galactic potential. The results of this analysis therefore show how these dynamical models can increase our knowledge of the potential.

The parameters of these potentials are given in the appendix, Table 4, and we refer to them in the text by their thin disc scalelengths.

We consider catalogues of “observations” of the discrete realisations. In general, a catalogue of N stars gives accurate values of the Galactic coordinates (b, l) , values of apparent magnitude m , colour $V - I$, and line-of-sight velocity v_{\parallel} that have moderate errors, and values of the parallax ϖ , proper motion $\boldsymbol{\mu}$, surface gravity g and metallicity Z that

**Figure 1.** Luminosity function given by equation (13), and used in all tests.

are probably significantly in error. We group the variables into two sets, the basic variables

$$\mathbf{u} \equiv (b, l, m, \varpi, \boldsymbol{\mu}, v_{\parallel}) \quad (11)$$

and additional astrophysical variables

$$\mathbf{s} \equiv (V - I, g, Z). \quad (12)$$

Note that \mathbf{u} has seven components, effectively a star’s phase-space coordinates (\mathbf{x}, \mathbf{v}) and its apparent magnitude m . For now we neglect interstellar extinction. Then a star’s absolute magnitude M is effectively specified by \mathbf{u} because its distance is fixed by \mathbf{x} .

As in Paper I, we restrict ourselves to the case of a single stellar population. This assumption ensures that there are no correlations between stellar type and kinematics: the distribution of stars in phase space is independent of their luminosities, colours, metallicities, etc. In this case we can confine discussion to the components of \mathbf{u} and neglect \mathbf{s} . We further assume that the luminosity function $F(M)$ is known to be

$$F(M) \propto \begin{cases} -14.9 + 21M - 5.4M^2 \\ \quad + 0.59M^3 - 0.019M^4 & \text{for } 1 < M < 19 \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

which is a simple polynomial approximation to the general V -band luminosity function described in Binney & Merrifield (1998), Table 3.16. This function is plotted in Figure 1 and satisfies the normalisation condition

$$1 = \int_{-\infty}^{\infty} dM F(M). \quad (14)$$

Throughout this paper we use the notation $F(m, s)$ to mean the luminosity function of an apparent magnitude m at a heliocentric distance s where, as we neglect extinction,¹

$$F(m, s) \equiv F(m - 5 \log(s/10 \text{ pc})). \quad (15)$$

We consider three catalogues of “observations” that give us information on the variables \mathbf{u}^α for each star α . In each case the catalogue contains $N_\alpha = 10\,000$ stars, all at Galactocentric latitude $b > 30^\circ$ and with apparent magnitude $m < 17$. We assume that no other selection effects bias the

¹ Here and throughout this paper we use log to mean \log_{10} .

$\Sigma_{d,0,\text{thin}}$	$R_{d,\text{thin}}$	$z_{d,\text{thin}}$	$\Sigma_{d,0,\text{thick}}$	$R_{d,\text{thick}}$	$z_{d,\text{thick}}$	$\rho_{b,0}$	$\rho_{h,0}$	r_h
$753.0 \text{ M}_\odot \text{ pc}^{-2}$	3 kpc	0.3 kpc	$182.0 \text{ M}_\odot \text{ pc}^{-2}$	3.5 kpc	0.9 kpc	$94.1 \text{ M}_\odot \text{ pc}^{-3}$	$0.0125 \text{ M}_\odot \text{ pc}^{-3}$	17 kpc

Table 1. Parameters of the Mc11 model used in the construction of the particle model. Note that the potentials with differing scaleheights used in this study have identical parameters except the two disc scaleheights (which are held with $z_{d,\text{thick}}/z_{d,\text{thin}} = 3$, with the quoted scaleheight always being that of the thin disc). The potentials with varying R_d have different parameters, listed in Table 4.

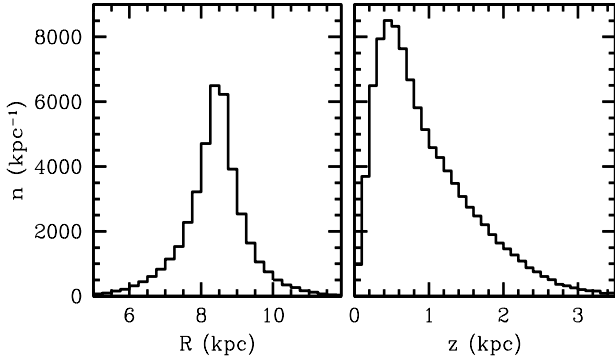


Figure 2. Histograms of the number of stars in the catalogue as a function of Galactocentric radius (left) and distance from the Galactic plane (right).

catalogue. For each observed star we assume that the quoted Galactic coordinates (b, l) are exact, as is the apparent magnitude m (this is equivalent to the statement that the uncertainty in m is much smaller than the scale on which $F(M)$ varies). We then have:

- A catalogue with measurements of parallax with uncertainty $\sigma_\varpi = 0.2 \text{ mas}$, measurements of line-of-sight velocity with uncertainty $\sigma_\parallel = 5 \text{ km s}^{-1}$, and measurements of proper motion with uncertainty (in each direction) of $\sigma_\mu = 0.2 \text{ mas yr}^{-1}$.
- A catalogue with parallax and proper motion measurements with the same uncertainty as previously, but with no measurement of the line-of-sight velocity.
- A catalogue with proper motion measurements with the same uncertainty as previously, but with no line-of-sight velocity or parallax measurements.

To convert values from Galactocentric coordinates to Heliocentric coordinates \mathbf{u} (to create this catalogue), or vice versa (to analyse the catalogue) we need to assume a position and velocity for the Sun. In all cases we assume that the Sun is at a Galactocentric radius $R_0 = 8.5 \text{ kpc}$ moving with the peculiar velocity found by Schönrich, Binney, & Dehnen (2010) with respect to the local standard of rest in the currently hypothesised potential.

In Figure 2 we show histograms of the number of stars in the catalogue as a function of Galactocentric R and z . We use the true positions of the stars to produce these histograms, rather than adding any uncertainties.

3 ASSESSING A MODEL LIKELIHOOD: TWO APPROACHES

3.1 General notation

Before describing our new method of solving the problem of determining model likelihoods, we explain our notation (which is very similar to that used in Paper I). We assume that the errors in the observed quantities are independent and can be modelled by Gaussian probability distributions

$$G(u, \bar{u}, \sigma) \equiv \frac{1}{\sqrt{2\pi}\sigma^2} e^{-(u-\bar{u})^2/2\sigma^2}. \quad (16)$$

We attach primes to the true values of measured quantities to distinguish them from the measured values. For brevity we use the notation

$$G_i^j(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) \equiv \prod_{k=i}^j G(u_k^\alpha, u'_k, \sigma_k^\alpha). \quad (17)$$

We find the likelihood \mathcal{L} of a model as a product over stars of the probabilities of measuring the values \mathbf{u}^α given the model:

$$\begin{aligned} \mathcal{L} &= \prod_\alpha \mathcal{L}_*^\alpha \equiv \prod_\alpha P(\mathbf{u}^\alpha | \text{Model}) \\ &= \prod_\alpha \int d^7 \mathbf{u}' G_1^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) \times P(\mathbf{u}' | \text{Model}). \end{aligned} \quad (18)$$

Any quantity that is not given in the catalogue (such as v_\parallel in two of our catalogues) can be considered to have a sufficiently large σ that the Gaussian density is effectively constant for all relevant values of the variable.

The probability $P(\mathbf{u}' | \text{Model}) d^7 \mathbf{u}'$ is the probability that a randomly chosen star in the catalogue has the true values \mathbf{u}' , for a particular model. Specifically $P(\mathbf{u}')$ is given by

$$P(\mathbf{u}' | \text{Model}) = A S(\mathbf{u}') F(M) f(\mathbf{x}, \mathbf{v}) \left| \frac{\partial(M, \mathbf{x}, \mathbf{v})}{\partial(\mathbf{u}')} \right|, \quad (19)$$

where the selection function $S(\mathbf{u}')$ is the probability that if a star with observables \mathbf{u}' exists it will be included in the catalogue, and the normalisation factor A depends on the DF, the luminosity function and the survey selection effects via the equation

$$1/A = \int d^3 \mathbf{x} d^3 \mathbf{v} f(\mathbf{x}, \mathbf{v}) \int dM S(\mathbf{x}, \mathbf{v}, M) F(M). \quad (20)$$

As we assume that the DF $f(\mathbf{x}, \mathbf{v})$ is that of an equilibrium dynamical model, we have

$$\int d^3 \mathbf{x} d^3 \mathbf{v} f(\mathbf{x}, \mathbf{v}) = (2\pi)^3 \int d^3 \mathbf{J} f(\mathbf{J}) = 1, \quad (21)$$

where the integrals are over all phase space and action space, respectively.

Our catalogue contains stars at Galactic latitudes $b > b_{\text{lim}}$ and apparent magnitude $m < m_{\text{lim}}$ with $b_{\text{lim}} = 30^\circ$ and

$m_{\text{lim}} = 17$. We assume that, other than these limits, the probability of a star entering the catalogue is independent of its properties, i.e.

$$S(\mathbf{u}') = \begin{cases} \text{const} & \text{for } b > b_{\text{lim}}, m < m_{\text{lim}} \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

Note that as this constant then appears in both the numerator and denominator (implicitly) of eq. (19), we can ignore it.

Since the potential Φ determines the relationship between $f(\mathbf{x}, \mathbf{v})$ and $f(\mathbf{J})$, the normalisation factor A (eq. 20) is a function of Φ in addition to $f(\mathbf{J})$, so we might write $P(\mathbf{u}'|\text{Model}) \equiv P(\mathbf{u}'|f, \Phi, F, S)$.

Formally, the best way to constrain the potential is to marginalise over all possible DFs (Magorrian 2006, 2013). While our use of DFs that are functions of \mathbf{J} opens up this exciting possibility, we do not attempt it in this study. Ting et al. (2012) marginalised over the parameters of a single pseudo-isothermal DF. This relatively low-dimensional marginalisation still required at least 1 to 2 orders of magnitude longer than simply finding the maximum likelihood for a DF of the assumed form in a given potential. They found that this marginalisation made little difference to their results (Ting, priv. comm.). In this study we only consider the maximum likelihood for a DF of the assumed form in the chosen Φ . This is very much like the approach used by Schwarzschild modellers.

3.2 Evaluating with $\mathbf{x}(\boldsymbol{\theta}, \mathbf{J})$

We first consider the approach to likelihood evaluation that was successfully employed in Paper I. This is founded on a library of tori, each of which yields $(\mathbf{x}(\boldsymbol{\theta}, \mathbf{J}), \mathbf{v}(\boldsymbol{\theta}, \mathbf{J}))$. The problems we encounter would be at least as serious if we were to replace the torus library by an orbit library of the type introduced by Schwarzschild (1979). Our description of the use of tori will be brief; a reader looking for more detail should read Sections 4 & 5 of Paper I.

A torus provides complete knowledge of the orbit with actions \mathbf{J} in a given potential. We therefore convert the integrals in equations (18) & (20) into integrals over $(\boldsymbol{\theta}, \mathbf{J})$. The normalisation factor A (eq 20) is rewritten as

$$1/A = \int f(\mathbf{J}) \phi(\mathbf{J}) d\mathbf{J}, \quad (23)$$

where

$$\phi(\mathbf{J}) \equiv \int d^3\boldsymbol{\theta} \int dm F(m, s) S(\mathbf{x}, \mathbf{v}, m). \quad (24)$$

Note that $\phi(\mathbf{J})$ depends on the potential. Regions of action space with $\phi(\mathbf{J}) = 0$ (in a given potential) do not contribute to the calculation and can be ignored – though the analytic forms we use for the DF do make predictions of the DF outside the survey volume, which can then be tested by further data.

As in Paper I we use the approximations that the measurements of b , l and m are exact to perform three of the seven integrals in equation (18), leaving integrals over \mathbf{J} and

s' (i.e. along the line of sight):

$$\begin{aligned} \mathcal{L}_*^\alpha &= A \int d^3\mathbf{J} f(\mathbf{J}) \int d^3\boldsymbol{\theta} dM G_1^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) F(M) S(\mathbf{u}') \\ &= A \int d^3\mathbf{J} \int ds' \left| \frac{\partial(\boldsymbol{\theta})}{\partial(b, l, s')} \right| \\ &\quad \times G_4^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) F(m, s') S(\mathbf{u}'), \end{aligned} \quad (25)$$

where $\mathbf{u}'(\boldsymbol{\theta}, \mathbf{J}, m)$ and $\boldsymbol{\theta}$ is now a function of s' because (b, l) are known.

The principle of Monte-Carlo integration is now invoked to convert these two integrals into sums over points \mathbf{J}_k that have been selected with a sampling density $f_S(\mathbf{J})$ that will be described below. Then we have

$$\mathcal{L}_*^\alpha = \sum_k \frac{f(\mathbf{J}_k)}{f_S(\mathbf{J}_k)} \text{LOSI}_{k,\alpha} / \sum_k \frac{f(\mathbf{J}_k)}{f_S(\mathbf{J}_k)} \phi_k, \quad (26)$$

where we have used the notation $\phi_k \equiv \phi(\mathbf{J}_k)$, and introduced the line-of-sight integral for a given \mathbf{J}_k and observation α

$$\text{LOSI}_{k,\alpha} = \int ds' \left| \frac{\partial(\boldsymbol{\theta})}{\partial(b, l, s')} \right|_k F(m, s') G_4^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) S(\mathbf{u}'). \quad (27)$$

The computation of these $N_k \times N_\alpha$ line-of-sight integrals dominates the computing budget for these calculations. Note that given our choice of selection function (eq. 22), we clearly have $S(\mathbf{u}') = \text{const} \neq 0$ in all cases. The Jacobian $\left| \frac{\partial(\boldsymbol{\theta})}{\partial(b, l, s')} \right|_k$ can be found using the torus machinery (Binney & McMillan 2011), and is closely related to the density of the orbit.

This approach allows us to reuse the values $\text{LOSI}_{k,\alpha}$ and ϕ_k that we have determined for the calculation of \mathcal{L} for many DFs in a given potential, but note that both depend critically on the relationship between \mathbf{u}' and $(\boldsymbol{\theta}, \mathbf{J})$, which depends on the potential, so they cannot be reused when we move to a new potential.

3.3 Evaluating with $\mathbf{J}(\mathbf{x}, \mathbf{v})$

Given that we have expressions for \mathbf{J} in terms of (\mathbf{x}, \mathbf{v}) , which we find using the Stäckel approximation, we can evaluate the normalisation factor A directly from equation (20) rather than its angle-action reformulation (23). We use the sampling density $f_S(\mathbf{x}, \mathbf{v})$ described below to ensure the points are concentrated where the integrand is largest. Then we have to evaluate the Monte-Carlo sum

$$1/A = \frac{1}{N} \sum_{k=1}^N \frac{f(\mathbf{J}(\mathbf{x}_k, \mathbf{v}_k))}{f_S(\mathbf{x}_k, \mathbf{v}_k)} \int dm F(m, r_k) S(\mathbf{x}_k, \mathbf{v}_k, m). \quad (28)$$

Since $1/A$ is a factor of every star's likelihood \mathcal{L}_*^α , the total likelihood $\mathcal{L} \propto A^{-N_*}$. Consequently, if we have a fractional error of δ in the value of A , the error in $\log \mathcal{L}$ is $\sim 0.43 N_* \delta$, so for our catalogues of 10 000 stars, we would require an uncertainty of ~ 0.02 per cent in A to limit the uncertainty in $\log \mathcal{L}$ to order unity. Numerical experiments to be described below reveal that with 4×10^6 sample points the uncertainty in A is ~ 0.4 per cent, $\sim 10^9$ sample points are required to determine the $\log \mathcal{L}$ to order unity. This is a very challenging requirement!

Fortunately the absolute value of \mathcal{L} is not important. What matters is the ratio of two given values of \mathcal{L} . We can minimise the numerical noise in this ratio by fixing the

points $(\mathbf{x}_k, \mathbf{v}_k)$ at which we evaluate the Monte-Carlo sum (eq 28) for all models, independently of the potential.

The great advantage of using $\mathbf{J}(\mathbf{x}, \mathbf{v})$ rather than $\mathbf{x}(\boldsymbol{\theta}, \mathbf{J})$ is that it is now possible to derive a Monte-Carlo sum for the likelihood in eq. (18) that runs over observable points \mathbf{u}' that are determined by the observations and, again, are independent of the potential. Moreover we will be able to arrange that each star's sample points will be concentrated within its error ellipsoid, rather than distributed regardless of where the star is observed.

We start by writing the integral associated with the α th star as

$$\mathcal{L}_*^\alpha = A \int d^7 \mathbf{u}' \left| \frac{\partial(\boldsymbol{\theta}, \mathbf{J}, M)}{\partial(\mathbf{u}')} \right| \times G_1^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) f(\mathbf{J}) F(M) S(\mathbf{u}'), \quad (29)$$

where the Jacobian is simply

$$\left| \frac{\partial(\boldsymbol{\theta}', \mathbf{J}', M')}{\partial(\mathbf{u}')} \right| = s^6 \cos b = \frac{\cos b}{\varpi^6}. \quad (30)$$

Then for each star we choose a sampling density $\xi(\mathbf{u}'|\mathbf{u}^\alpha)$ that causes the sample points to be concentrated in the region of \mathbf{u}' space that dominates the integral. In the case of small errors, this region is the inner $\sim 3\sigma$ of the error ellipsoid. In the case of serious errors – for example if a variable such as v_{\parallel} has not been measured or the measured value of ϖ is negative – this region is a subset of the error ellipsoid that is consistent with reasonable expectations of how stars are distributed in phase space. Again assuming that the measured values of l^α, b^α and m^α are exact, we are led to the sampling density

$$\xi(\mathbf{u}'|\mathbf{u}^\alpha) = \begin{cases} 0 & \text{for } (l, b) \neq (l^\alpha, b^\alpha) \\ C^\alpha \left| \frac{\partial(\boldsymbol{\theta}', \mathbf{J}', M')}{\partial(\mathbf{u}')} \right| G_4^7(\mathbf{u}^\alpha, \mathbf{u}', \boldsymbol{\sigma}^\alpha) & \\ \times f_S(\mathbf{x}', \mathbf{v}') F(m^\alpha, s) & \text{otherwise,} \end{cases} \quad (31)$$

where C^α is the normalising constant. With this sampling density the integral (eq. 29) reduces to the Monte-Carlo sum

$$\mathcal{L}_*^\alpha = \frac{A}{N C^\alpha} \sum_{k=1}^N \frac{f(\mathbf{J}'_k)}{f_S(\mathbf{x}'_k, \mathbf{v}'_k)}, \quad (32)$$

where $(\mathbf{x}'_k, \mathbf{v}'_k)$ is determined from \mathbf{u}'_k , and $\mathbf{J}'_k = \mathbf{J}(\mathbf{x}'_k, \mathbf{v}'_k)$. We note that C^α is independent of $f(\mathbf{J})$ and Φ , so it does not vary as we explore different DFs and potentials. Since we are only ever interested in the ratio between likelihoods calculated for different DFs and potentials, we do not need to compute C^α .

3.4 Choice of sampling density

A good choice of the sampling density $f_S(\mathbf{x}, \mathbf{v})$ used in these calculations reduces numerical noise by making the individual contributions to the Monte-Carlo sums as nearly equal as possible. We achieve this goal by choosing $f_S(\mathbf{x}, \mathbf{v})$ to be a good guess at the phase-space distribution of the population that our catalogue samples. It is important to ensure that there are no points with $f_S \ll f$, as these will then dominate the integrals (eqs. 28 & 32), making them very noisy. We have based these on a product of a double-exponential density in real space with a triaxial Gaussian velocity distribution. The principal axes of the velocity distribution are

aligned with the R, z and ϕ directions, with dispersions that vary in proportion to $\exp(-R/8 \text{ kpc})$. The means of the v_R and v_z components are zero, while the mean of v_ϕ is

$$\langle v_\phi \rangle = v_c - v_a(R) \quad (33)$$

with constant $v_c = 245 \text{ km s}^{-1}$, and asymmetric drift velocity $v_a(R) \propto \sigma^2$.

We use the sum of two such discs (approximating the thin and thick discs) as f_S . We set the dispersions at the Solar radius $\sigma_{R,\odot} = \sigma_{z,\odot} = 30 \text{ km s}^{-1}$, $\sigma_{\phi,\odot} = 40 \text{ km s}^{-1}$, and $v_{a,\odot} = 15 \text{ km s}^{-1}$ for a disc with scaleheight 0.3 kpc; and $\sigma_{R,\odot} = \sigma_{z,\odot} = 50 \text{ km s}^{-1}$, $\sigma_{\phi,\odot} = 60 \text{ km s}^{-1}$, and $v_{a,\odot} = 30 \text{ km s}^{-1}$ for a disc with scaleheight 1 kpc. 30 per cent of f_S is associated with the thick disc component, and the rest with the thin disc.

4 THE PROBLEM WITH ORBIT-BASED METHODS

To understand why orbit-based methods are extremely ill-suited to determining the likelihood of a stellar catalogue, we need to look at the numerator of equation (26). The main calculation for each star α is finding the line-of-sight integral $\text{LOSI}_{k,\alpha}$ for each orbit used in the Monte-Carlo integration. These are then summed (with some weights) to find the probability that this star is predicted by the model. $\text{LOSI}_{k,\alpha}$ is essentially the integral down the line-of-sight of the probability of a given orbit giving the observed value of \mathbf{u} . With increasingly precise data, the number of orbits for which this probability is non-negligible anywhere along the line of sight diminishes.

The computation of a star's likelihood proceeds by considering each point along the line of sight to the star, and then for each orbit finding the velocities that a star on that orbit will have at that point. Each such velocity then contributes to the probability a factor proportional to

$$\exp \left[- \left(\frac{(v_{\parallel}^\alpha - v_{\parallel})^2}{2\sigma_{\parallel}^2} + \frac{|\boldsymbol{\mu}^\alpha - \boldsymbol{\mu}|^2}{2\sigma_{\mu}^2} \right) \right]. \quad (34)$$

As we proceed along the line of sight, the observables $(v_{\parallel}, \boldsymbol{\mu})$ predicted by the orbit gradually change, so the overall contribution of the orbit to the likelihood comes from a line drawn through the error ellipsoid. The contribution is large or small depending on how close the line comes to the centre of the ellipsoid.

When we change potential, our orbits necessarily change and the lines through the error ellipsoid change. It may happen that our new orbit library has an orbit that yields a line that comes very close to the centre of the ellipsoid, whereas the old orbit library jumped from an orbit that gave a line passing closest to the centre say 2σ on one side to an orbit that passes closest 2σ on the other side of the ellipsoid's centre. This old library provided no orbit that makes the given star very probable even though, with a denser sampling of action space, such an orbit would have arisen. Hence with the old library rather than the new, the star is declared improbable even though it is in fact probable.

A subsidiary issue is that with the new library some of the values \mathbf{J}_k which had $\phi(\mathbf{J}_k) = 0$ in the original potential (i.e. orbits that do not enter the survey volume, and were

thus irrelevant to our calculation), have $\phi(\mathbf{J}_k) \neq 0$ in the new library and have to be considered, and vice versa.

An indication of how significant these discreteness effects are is that changing the potential from one with scaleheight 300 pc to one with scaleheight 350 pc causes the calculated values $\text{LOSI}_{k,\alpha}$ to change by on average a factor ~ 10 .

4.1 Extent of the problem

To see the impact of this problem, we now turn to our three catalogues described in Section 2.4. Numerical experiments show that keeping the same values of \mathbf{J}_k as we change potentials is little better than taking an entirely new Monte-Carlo sampling of \mathbf{J}_k , so we now explore the case in which we keep the same Mc11 potential, Φ , and the same catalogue of stars, but use a new sample of actions \mathbf{J}_k to calculate the likelihoods (eq. 26). In each case we use the DF f_{true} that was used to create the catalogue as both the sampling density and the DF being tested in equation (26).

In principle it would be ideal to take a large number of Monte-Carlo samples, each containing N_k values of \mathbf{J} , determine the likelihood in each, and directly determine the scatter. However this process is prohibitively expensive computationally, so we use a quicker alternative. We calculate the integrals ϕ_k and $\text{LOSI}_{k,\alpha}$ (equations 24 & 27 respectively) for 100 000 values of \mathbf{J}_k . Then we choose at random subsets of 12 500 or 25 000 or 50 000 values of \mathbf{J}_k and use these values to calculate for each star α the standard deviation $\sigma(\log \mathcal{L}_*^\alpha)$ of the resulting values of $\log \mathcal{L}_*^\alpha$.

Fig. 3 shows the distributions of these standard deviations for each library size and each quality of data. Increasing the quality of the data broadens the distribution of $\sigma(\log \mathcal{L}_*^\alpha)$ and shifts its mean to higher values. Even for $N_k = 50\,000$ all stars have non-negligible values of $\sigma(\log \mathcal{L}_*^\alpha)$ and only in the case of the highest-quality data do the distributions have a heavy tail to large values. Thus the problem is not so much the existence of a few outliers but that our Monte-Carlo sums give rise to excessive uncertainty in the likelihoods of all stars.

Note that the largest torus libraries considered are more than an order of magnitude larger than the orbit libraries used for typical Schwarzschild modelling of external galaxies (e.g. van den Bosch et al. 2008), though we do not “dither” orbits, which is an approach that can significantly increase the effective resolution of Schwarzschild models.

From Fig. 3 it is evident that $\sigma(\log \mathcal{L}_*^\alpha)$ can be beaten down by increasing the number N_k of orbits in one’s library. In Paper I we used the Shannon entropy to measure the extent to which an individual observation \mathbf{u}^α is probed by a given library of tori. The entropy is

$$S^\alpha = - \sum_k^N p_k^\alpha \ln p_k^\alpha, \quad (35)$$

where p_k^α is the fraction of the calculated likelihood \mathcal{L}_*^α contributed by the k th torus. This is

$$p_k^\alpha \equiv \frac{\text{LOSI}_{k,\alpha}}{\sum_k \text{LOSI}_{k,\alpha}}. \quad (36)$$

Clearly $S^\alpha = 0$ if there is only one contribution, and $S^\alpha = \ln N$ if N tori provide equal contributions to the star’s

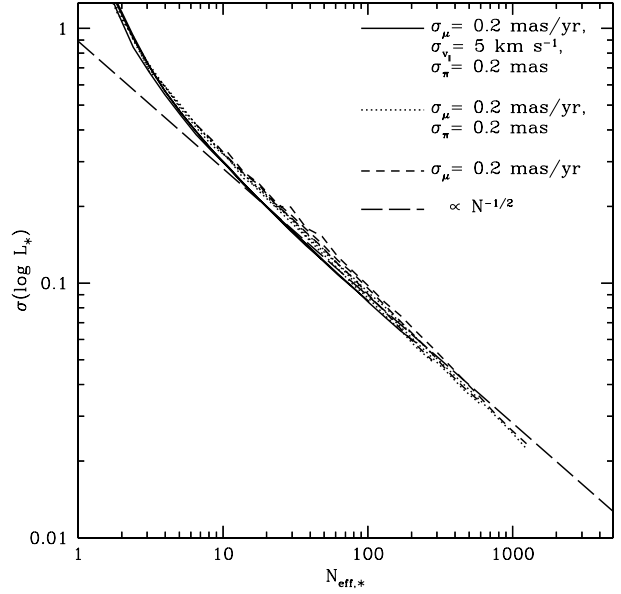


Figure 4. Torus modelling: standard deviations of a stars’ likelihood, $\sigma(\mathcal{L}_*)$ as a function of the effective number of tori contributing to the calculation of each likelihood ($N_{\text{eff},*}$ eq. 37). The three star catalogues, which differ in the completeness of their data, are represented by solid, short-dashed or dotted lines. There are three lines for each catalogue in the figure, corresponding to three different sizes of torus library, $N_k = 12500, 25000$ and 50000 . The long-dashed line is $\sigma(\mathcal{L}_*) = 0.9 N_{\text{eff},*}^{-1/2}$, which fits all points with $N_{\text{eff},*} \gtrsim 10$ quite well.

probability. We can therefore define an effective number of contributing values \mathbf{J}_k ,

$$N_{\text{eff},*}^\alpha \equiv \exp(S^\alpha), \quad (37)$$

which is the number of tori with equal p_k^α that would give an entropy S^α . Clearly we expect that, for a given library size N_k , the typical value of $N_{\text{eff},*}^\alpha$ will become smaller as the observational data become more precise.

Figure 4 plots the standard deviation $\sigma(\log \mathcal{L}_*^\alpha)$ against the effective number $N_{\text{eff},*}^\alpha$ derived from the average of the entropies S^α over a sample of equivalent torus libraries. We see that all data points lie close to a universal relation, that is independent of library size or data quality. Moreover, this relation asymptotes to the relation $\sigma(\log \mathcal{L}_*^\alpha) \propto \sqrt{N_{\text{eff},*}^\alpha}$. This result conclusively proves that the scatter in likelihood values is generated by Poisson noise, and enables us to predict how large $N_{\text{eff},*}^\alpha$ needs to be to beat the noise down to any given level for any data set.

Figure 5 shows how the errors in individual star likelihoods combine to produce errors in model likelihoods \mathcal{L} by showing histograms of the offsets $\Delta(\log \mathcal{L})$ between $\log \mathcal{L}$ and its mean over all torus libraries. These distributions are roughly Gaussian so their widths are characterised by the standard deviation $\sigma(\log \mathcal{L})$ of $\log \mathcal{L}$.

Figure 6 shows $\sigma(\log \mathcal{L})$ as a function of torus library size N_k for all three catalogues. The more precise the data, the larger $\sigma(\log \mathcal{L})$ is, but in each case $\sigma(\log \mathcal{L})$ declines approximately in proportion to $\sqrt{N_k}$, but is still ~ 10 even for our largest library and our lowest quality data.

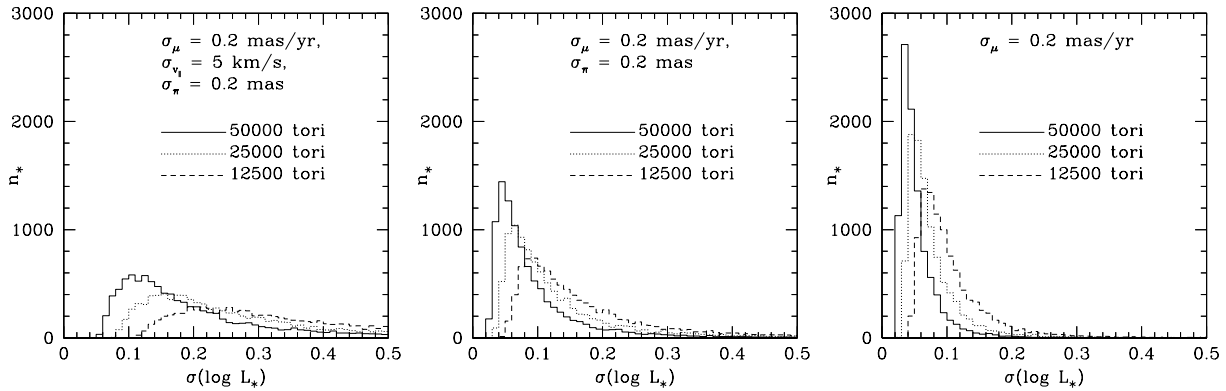


Figure 3. Torus modelling: histograms of the standard deviations $\sigma(\log L_*)$ calculated from torus libraries of various sizes (as labelled) for each of the 10 000 stars in each of our three catalogues (also as labelled).

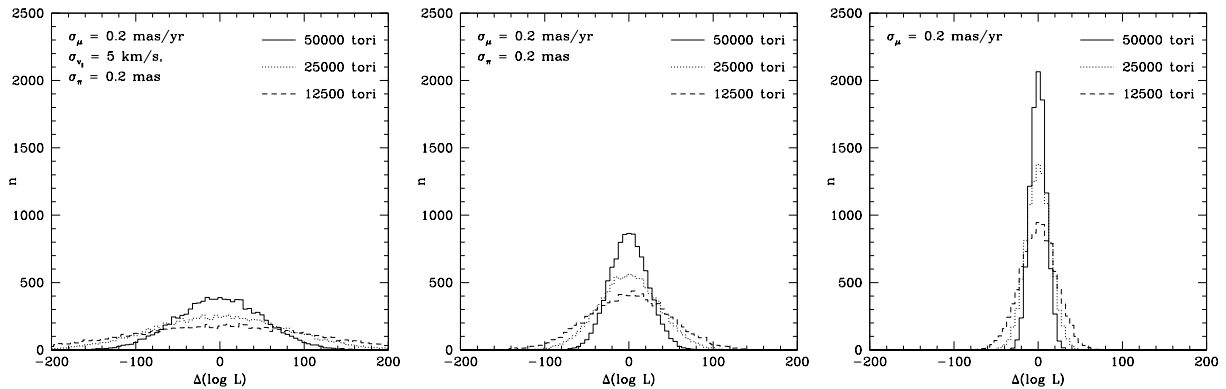


Figure 5. Torus modelling: histograms of the value of $\log \mathcal{L}$ calculated for each of our three catalogues. Each value of $\log \mathcal{L}$ is computed 10 000 times using a different torus library. Distributions are shown for three sizes of torus library: $N_k = 12\,500, 25\,000$ or $50\,000$ tori. From each value of $\log \mathcal{L}$ we have subtracted the mean value of its set.

4.1.1 Implications for determination of Φ

The key question is: will this uncertainty on the value of \mathcal{L} prevent us from making useful inferences about the Galactic potential? To answer this question we now use tori to determine the likelihoods of our three catalogues given various choices of potential. In Figure 7 we show the peak values of $\log \mathcal{L}$ calculated for models constrained to have two pseudo-isothermal discs in potentials with varying disc scalelengths (upper row) and scaleheights (lower row) found from a Monte-Carlo sum over 100 000 values \mathbf{J}_k (i.e. 100 000 tori). The true potential has disc parameters $(R_d, z_d) = (3, 0.3)$ kpc. We show (approximate) error bars on each value – these are found by extrapolating the relationship shown in Figure 6 to $N_k = 100\,000$. Note, therefore, that this is the uncertainty on \mathcal{L} due to the imperfect analysis, and is not intrinsic to the observational data.

In each case, the models with $R_d = 2.5$ can clearly be ruled out, but the models with $R_d = 3.5$ have the same \mathcal{L} as the true potential to within the error bars. The results when varying z_d are even less encouraging, as all the calculated values of \mathcal{L} agree to within the error bars, and the incorrect potentials ($z_d = 0.25$ or 0.35) are sometimes preferred to the true potential. Clearly the true difference in $\log \mathcal{L}$ is significantly smaller than the uncertainty.

These calculations would require over a week on a single desktop cpu for each potential (though they are simple to parallelise). The majority of this time is spent calculating the line-of-sight integrals $\text{LOSI}_{k,\alpha}$.

4.2 Why could we determine the DF?

Our aim in this section is to explain clearly why the torus-based method of Paper I could determine the DF to good precision but fails here when extended to determination of Φ . To this end we consider the toy problem illustrated in Figure 8. We use the Monte-Carlo principle to estimate the ratio of two integrals I_i of functions $\eta_i(x)$ of one variable that we know to be products of the Gaussian of unit dispersion and zero mean and functions $f_i(x)$ that vary on scales larger than unity. Thus

$$I_i = \int \eta_i(x) dx = \int dx, f_i(x) \times G(x, 0, 1), \quad (38)$$

where for the f_i we adopt

$$\begin{aligned} f_1(x) &= 1 + 0.01x \\ f_2(x) &= 0.95 + 0.03x. \end{aligned} \quad (39)$$

With these choices the true values are $I_1 = 1$ and $I_2 = 0.95$. The Gaussian represents a star's error ellipsoid and

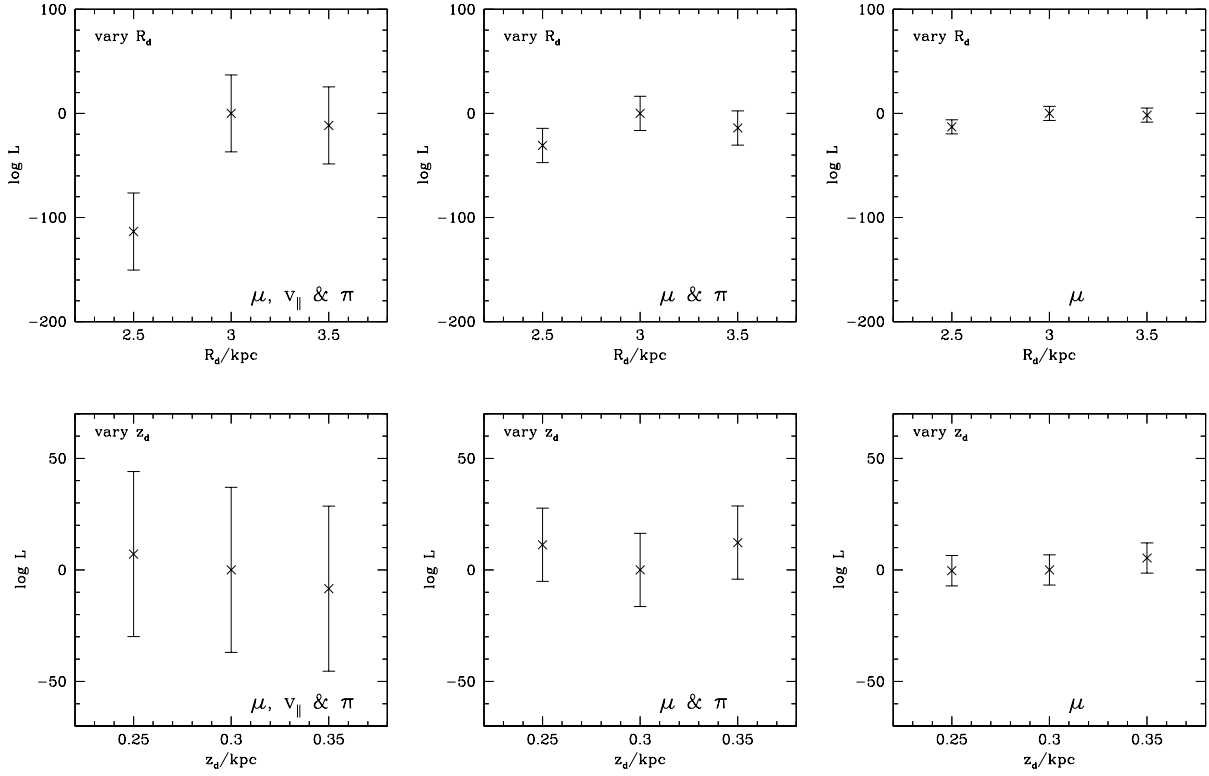


Figure 7. Torus modelling: differences in the largest values of log likelihood obtained by varying the DF in a given Φ and the value for the true Φ . The DF $f(\mathbf{J})$ is assumed to be of the form (2). In the upper row the scalelength R_d of the disc that contributes to Φ is varied, while in the lower row the its scaleheight z_d is varied. The true values are $(R_d, z_d) = (3, 0.3)$ kpc. The computations use $N_k = 100\,000$ tori. The error bars are approximate, and found from an extrapolation to $N_k = 100\,000$ of the relations shown in Fig. 6, assuming that $\sigma \log \mathcal{L} \propto N_k^{-1/2}$.

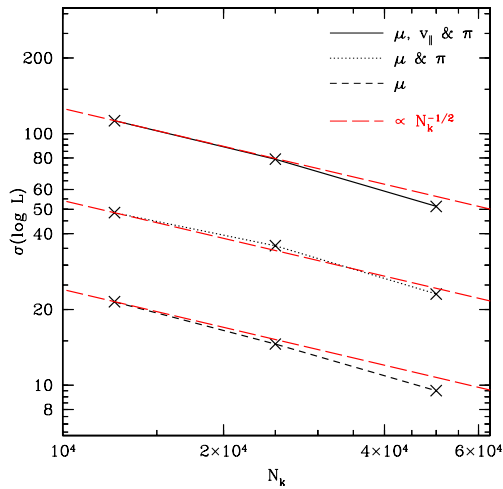


Figure 6. Torus modelling: the standard deviations of the histograms plotted in Fig. 5 as functions of the number N_k of tori employed. The red long-dashed lines have slope $-1/2$ so we see that the uncertainty in $\log \mathcal{L}$ declines approximately as $N_k^{1/2}$.

the f_i represent candidate DFs. Figure 8 shows the process of Monte-Carlo evaluation of the I_i . In the top left panel I_1 is found to be $I_1 = 1.129$ rather than its true value, unity. In

the top right panel independently sampled points are used to estimate that $I_2 = 1.206$, so the ratio of the integrals is $I_2/I_1 = 1.068$ rather than 0.95 as it should be. The lower panels show re-evaluations of I_2 using points that are not independent of those used to evaluate I_1 : in the bottom-right panel we use exactly the same points to find $I_2/I_1 = 0.953$, within 0.4% of its true value, while in the bottom left panel we use points that are shifted by 1.5 to the left, and find $I_1/I_2 = 0.860$.

This experiment shows that if we use the same sampling points we can determine the ratio of two integrals like those we have to evaluate to obtain \mathcal{L}_*^c much more accurately than we can determine either integral individually because the Poisson noise in the evaluation largely cancels from the ratio. If we use different sampling points to evaluate each integral, the Poisson noise does not cancel and the ratio is even less accurate than the individual integrals. In Paper I we used one set of sampling points to evaluate the likelihood of every DF, so the Poisson noise made little contribution to the differences of the log likelihoods of the DFs considered, and we could identify the true DF accurately. In Section 4.1 we were obliged to vary the sampling points between potentials and the Poisson noise in the differences of log likelihoods degraded performance to an unacceptable extent.

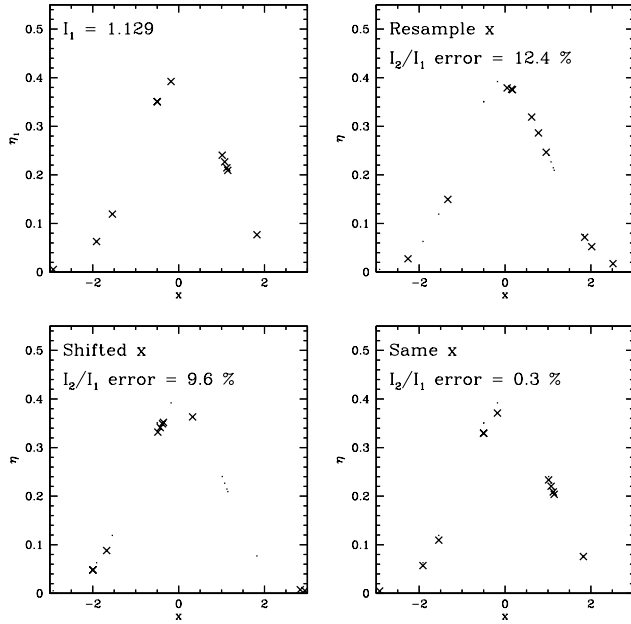


Figure 8. Toy example analogous to comparing the likelihoods of given data in different potentials. In each case we determine an integral I_i (eq. 38) by Monte-Carlo summation over 20 points taken randomly from the range $-5 < x < 5$. In the top-left panel we show the function values η_1 used to find I_1 (in the other panel these values are shown as dots, for comparison). The sum yields $I_1 = 1.129$ rather than unity. In the other three panels we show the summed values of η_2 when we have either completely resampled the values of x (top right), or have shifted all the values by $\Delta_x = -1.5$ (with any points that fall below $x = -5$ replaced by ones at $x > 3.5$, bottom left), or have used the same values of x as were used for I_1 (bottom right). Although these last points do not yield a particularly accurate value of I_2 they do yield an accurate value for the ratio I_1/I_2 .

5 THE SOLUTION: USE $\mathbf{J}(\mathbf{x}, \mathbf{v})$

In Section 3.3 we explained how, if we have a way to find $\mathbf{J}(\mathbf{x}, \mathbf{v})$, the \mathcal{L}_*^α can be evaluated using samples of points \mathbf{u}'_k that are chosen specifically for each star in the catalogue and are never varied. We also explained that a second sample of points should be used to evaluate the normalising constant A for all potentials considered. We now show that this approach dramatically reduces the numerical noise that was so prominent in Section 4.1.

In summary, the scheme is:

- (i) For the calculation of A , sample N_A points (\mathbf{x}, \mathbf{v}) from the sampling density $f_S(\mathbf{x}, \mathbf{v})$ (ignoring any that lie outside the survey volume).
- (ii) Sample $N_{\mathbf{u}'}$ points for each star from the sampling density $\xi(\mathbf{u}'|\mathbf{u}^\alpha)$, (eq. 31).
- (iii) Choose a gravitational potential Φ , and determine \mathbf{J} for each of the N_A points used to determine A and each of the $N_\alpha \times N_{\mathbf{u}'}$ points used to then determine \mathcal{L} .
- (iv) Maximise \mathcal{L} in this potential by varying the parameters of the DF $f(\mathbf{J})$.
- (v) Return to step (iii), choosing a new potential.

This process is orders of magnitude faster than the torus approach, so we are able to carry out many more tests.

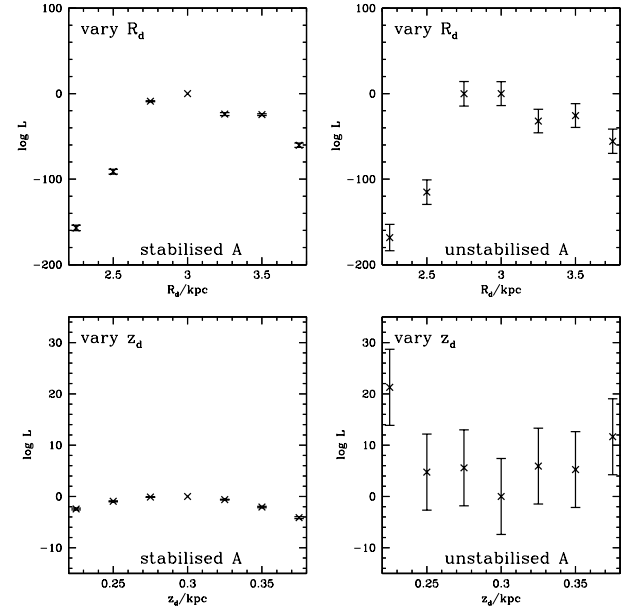


Figure 9. Using $\mathbf{J}(\mathbf{x}, \mathbf{v})$ when the data are error-free: Differences between the largest value of $\log \mathcal{L}$ obtained for a candidate Φ and the value obtained for the true Φ as we vary either the scale-length of the potential-generating disc (upper) or its scaleheight (lower). The left panels show results obtained using the fixed sets of sampling phase-space points throughout, while the right column shows the effect of choosing new points for the estimation of A for each trial Φ . Note that the range of $\log \mathcal{L}$ on the y-axis is less than half that in Fig. 7.

As a proof of principle, we show in the left column of Fig. 9 results for the case in which we have perfect observational data with the consequence that the Monte-Carlo sum for \mathcal{L}_*^α (eq. 31) requires just one point. These tests are very similar to those of Ting et al. (2012), except we have significantly a more complicated (and realistic) selection function that requires a more careful Monte-Carlo integration, and we do not marginalise over the parameters of our (somewhat more complicated) DF. We use $N_A = 4 \times 10^6$ points for our normalisation calculation (as opposed to 10^5 used by Ting et al: priv. comm.). Note that the equivalent calculation is essentially impossible to perform correctly with an orbit library, as the probability of an orbit in the library passing precisely through the observed phase-space location of a star is zero.

The top left panel of Fig. 9 shows the effect of systematically varying the scalelength R_d of the disc that contributes to Φ around its true value $R_d = 3$ kpc, while the lower left panel shows the effect of systematically varying the potential's scaleheight around its true value $z_d = 0.3$ kpc. The data points now reveal both R_d and z_d to good precision.

The right panels of Fig. 7 show the importance of preventing the Poisson noise in our estimate of A from scattering the data points by showing the points one obtains when new sampling points (\mathbf{x}, \mathbf{v}) are chosen for each trial Φ . The noise has a totally devastating impact on our ability to deduce z_d .

When we resample for each potential, we can determine the uncertainty in log likelihood simply by repeating the experiment several times for the same data set and determin-

Table 3. Uncertainties in z_d . The *intrinsic* uncertainty is the uncertainty due to the finite size and observational accuracy of the catalogue. We find this value by fitting a Gaussian in z_d to \mathcal{L} . The numerical uncertainty is the uncertainty introduced by the limited numerical precision of the integrals used to find \mathcal{L} .

Data	Best fit z_d	Intrinsic uncertainty	Numerical uncertainty
Exact	0.292	0.019	0.002
μ , v_{\parallel} & ϖ	0.294	0.021	0.006
μ & ϖ	0.272	0.032	0.012
μ	0.313	0.036	0.017

ing the standard deviation of the recovered values. When we do not resample, this approach is inadequate – the expected improvement is in the accuracy of the *relative* log likelihoods found. We therefore determine error bars by fixing the log likelihood found for the (known) true potential as zero in each experiment and finding the scatter in the relative value found in each given potential.

Figure 10 shows results obtained when we use $\mathbf{J}(\mathbf{x}, \mathbf{v})$ and fixed phase-space sampling points to analyse our three catalogues of varying completeness. The improvement over the results shown in Fig. 7 is dramatic – note that the scale in $\log \mathcal{L}$ for plots with varying z_d is an order of magnitude larger in Figure 10 than in Figure 7. The uncertainty in the difference between the $\log \mathcal{L}$ values of potentials that differ by 25 pc in their values of z_d is now as small as ~ 0.1 . With this level of uncertainty in $\log \mathcal{L}$ differences, it becomes possible to constrain the value of z_d strongly. Extrapolating the trend shown in Figure 6 suggests that achieving the same precision with tori would require the use of $\sim 10^9$ to 10^{10} tori for the three catalogues we consider.

We can quantify both the intrinsic uncertainty in Φ associated with our catalogues and the additional uncertainty produced by the Monte-Carlo integration. For the varying scaleheights (Fig. 10 lower panels) we can fit the values of $\log \mathcal{L}$ to a quadratic in z_d (i.e. approximate \mathcal{L} as Gaussian in z_d). We can then read off the most likely z_d , and its uncertainty σ_{z_d} , which is very close to the intrinsic uncertainty. If we do this many times (for many different Monte-Carlo sums) we can compare the most likely z_d found in each case and find the scatter in these values, which is the uncertainty associated with the Monte-Carlo integration.

Table 3 gives these uncertainties. With only 10 000 stars we can determine z_d with intrinsic uncertainty of less than 36 pc with only proper motion data, or less than 20 pc when line-of-sight velocities and parallaxes are also available. In each case the uncertainty introduced by the Monte-Carlo sums is significant smaller than the intrinsic uncertainty. Since the volumes of the error ellipsoids increases as the completeness decreases, to achieve a given precision more points are required in the Monte-Carlo sums (larger $N_{\mathbf{u}'}$) when the data are incomplete than when they are complete.

For varying scalelengths, it's clear that in the range analysed, \mathcal{L} is not well approximated by a Gaussian in R_d . The range in $\log \mathcal{L}$ found is much larger than in the scale-height case. With this analysis we can therefore only reasonably say that the uncertainty in R_d is significantly smaller than 250 pc.

Determining one of these likelihoods requires the calculation of $\sim 10^7$ values of the actions in a given potential, a process which takes ~ 10 minutes on an ordinary desktop cpu, and can easily be parallelised. This is $\sim 10^4$ times faster than the calculations using tori of Section 4.1. In fact, to achieve with tori the same precision we have achieved using $\mathbf{J}(\mathbf{x}, \mathbf{v})$ would demand $\sim 10^8$ more cpu cycles than were used for this section!

6 DISCUSSION

There are three different kinds of uncertainty associated with this work

- Irreducible statistical uncertainty. This is the uncertainty associated with the limited number of stars in the catalogue and their non-zero measurement errors.
- Numerical noise associated with the limited accuracy with which we evaluate integrals over the DF.
- Systematic errors associated with (i) inaccuracies in the transformations between angle-action and ordinary phase-space coordinates, and (ii) the use of particular functional forms for the DFs and potentials that we fit to the data.

We have found that the key to reducing the numerical noise to the point where it is possible to successfully determine the Galactic potential from a star catalogue is (i) evaluation of the DF at points that are fixed in the space of observables \mathbf{u} , and (ii) clustering these points within the error ellipsoid of each observed star, so we are sure to evaluate the DF throughout the region of phase space where each star might lie.

Given that the initial conditions of an orbit can be considered its integrals of motion, it might be argued that our prescription is readily implemented within the context of Schwarzschild modelling: we build our orbit library by integrating orbits from initial conditions that strategically cover each star's error ellipsoid.

One way to see the fatal weakness of this idea is *reductio ad absurdum*: we consider the limiting case of perfect data. Then only one orbit will be required to cover each star's error ellipsoid, and when we assign unit weight to each orbit, we will obtain perfect agreement with the data regardless of what potential we choose because the orbit started for one star has zero probability of being sampled at the location of another star. The potential can be constrained only to the extent that each orbit contributes non-trivially to the likelihood of more than one star.

Chanamé et al. (2008) achieve this goal by explicit binning of the model on the sky. This binning operation is essentially the means by which the density in the space of observables is constructed out of the otherwise uninformative orbital weights. The usefulness of binning decreases rapidly as the dimension d of the space or the data increases. We are currently considering the case $d = 6$, but once we include crucial spectral information in our models, d rises to $d = 10$ and beyond (e.g. Binney 2011).

An alternative approach to the problem of determining the Galactic potential is to use Jeans' equations (e.g. Binney & Tremaine 2008, §4.8) to relate gradients in the density and velocity dispersion of a suitable tracer population to the gradient in the potential. Recent studies using this method

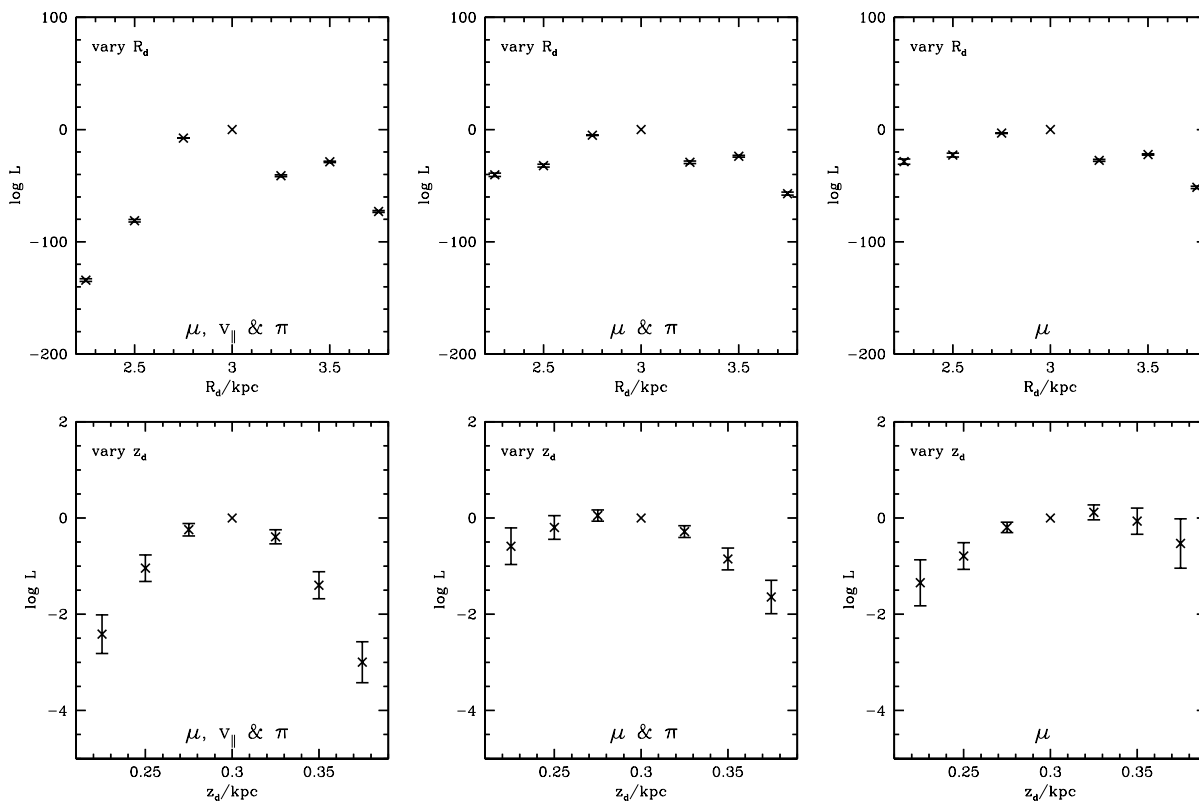


Figure 10. Results obtained with $\mathbf{J}(\mathbf{x}, \mathbf{v})$ with data of varying completeness. For each of our three catalogues studied in Fig. 7 we plot the largest value of $\log \mathcal{L}$ minus that obtained for the true potential as either the scale length of the disc that contributes to Φ is systematically varied (upper row) or the disc’s scaleheight z_d is systematically varied (lower row, note that the range of $\log \mathcal{L}$ on the y-axis is an order of magnitude smaller than in the equivalent plots in Figs 9 & 7). The number of points used for each star was $N_{u'} = 1000$ for the catalogue with measured μ, v_{\parallel} and ϖ and $N_{u'} = 2000$ points for the two catalogues with less complete data.

include Garbari et al. (2012) and Bovy & Tremaine (2012) – though it should be noted in the latter case that the velocity dispersions assumed were biased by up to a factor of 2 and had materially under-estimated errors (Sanders 2012b), so the quoted results will also be biased and offer spurious precision. Since this approach relies on the *gradient* in density, it is particularly susceptible to errors in the density profile, which become more likely for survey data with complicated selection effects, as the selection criteria are typically magnitude and colour, and vary with position on the sky.

Our ability to diagnose Φ depends crucially on components of our DF contributing to the likelihood of more than one star (e.g. Magorrian 2013). When $\mathbf{J}(\mathbf{x}, \mathbf{v})$ is available, sampling the error ellipsoids of stars works because our DF $f(\mathbf{J})$ is conjectured from the outset rather than estimated by binning products of weights and orbital probabilities. Because we require f to be a smooth function of \mathbf{J} , a change in the value of f at the actions of one star changes the value of f at the actions of many other stars in a way that depends on Φ . It is this principle that provides diagnostic power.

Our choice of parametrised form for $f(\mathbf{J})$ is therefore crucial. An excessively flexible form will simply fit the noise in the data. A badly chosen or insufficiently flexible form will produce biased results. For example, if we perform the tests with varying z_d as in Section 5, except that we only allow f to consist of a single quasi-isothermal disc (as opposed to the two discs it actually comprises), we are strongly biased

towards low values of z_d . DFs of the type used here have been shown to provide good fits to observational data in the Solar neighbourhood Binney (2010, 2012b), but it is clear that one must be careful not to over-constrain them at the expense of biasing estimates of Φ .

As Magorrian (2006, 2013) has stressed, Φ should really be found by marginalising over the DF rather than by finding the pair (f, Φ) that maximises the likelihood of the data. In statistical problems we often take the shortcut of seeking the most likely value of some variable rather than the variable’s expectation value, but the justification for this step has to be that the probability distribution is so sharply peaked around the most likely value that these two values are effectively indistinguishable. The classic example of how misleading this assumption can be, is provided by the thermal equilibrium of a macroscopic object, such as a diamond of N atoms. Since the probability that any of the diamond’s normal modes is in its i th excited state of energy E_i is proportional to the Boltzmann factor $\exp(-E_i/k_B T)$, the diamond’s ground state, in which all normal modes are unexcited, is by far the most probable state regardless of the temperature T . Yet a real diamond has negligible probability of being in its ground state: it is certain to be in a state that is higher in energy by $\sim 3Nk_B T$. The actual state is extraordinarily improbable, but there are so many states like it, that we can be certain the diamond is in *one* of them and not its enormously more probable ground state.

Thus it is dangerous to suppose, as we have done, that the Galaxy’s potential is the member of the pair (f, Φ) that has the highest probability: this may be a singular pair and nearly all the probability is associated with materially different pairs (f', Φ') , so these pairs would dominate the expectation of Φ if we marginalised over the DF. The key to this marginalisation is knowing how to sample the space of all possible DFs. Magorrian (2013) explains how this should be done, but we do not yet know whether doing so materially changes our conclusion regarding the form of Φ . Ting et al. (2012) marginalised over the parameters of their single pseudo-isothermal DF, and found that this made little improvement to their results as compared to simply finding a maximum likelihood. This is, however, still tied to the parametrised form of the pseudo-isothermal DF, and therefore does not really answer the question.

It is encouraging that the formulae for $\mathbf{J}(\mathbf{x}, \mathbf{v})$ introduced by Binney (2012a) are accurate enough to perform the analysis in Section 5 without biasing the results on the investigated scales. However, they are neither as general nor as accurate as the principle of torus construction – the latter is a systematic approximation scheme whose accuracy can be ramped up at will. Binney’s formulae are by contrast fixed: their accuracy cannot be systematically increased. They were introduced and validated in the context of the orbits of disc stars in the solar neighbourhood, and it not entirely clear why they work as well as they do for these orbits. Work needs to be done to optimise the extension of these formulae to the orbits of bulge and halo stars. Sadly, there is scant prospect that these formulae can be extended to the rotating non-axisymmetric potential of the Galactic bar, so the conclusion that our ability to diagnose Φ hangs by the slender thread of these formulae is a worrisome one.

Fortunately, values $\mathbf{J}(\mathbf{x}, \mathbf{v})$ can be obtained from tori: given a trial potential and a point (\mathbf{x}, \mathbf{v}) we estimate $(\boldsymbol{\theta}, \mathbf{J})$, perhaps from Binney’s formulae, and construct a trial torus. Then as described in McMillan & Binney (2008) we iteratively adjust \mathbf{J} until we obtain a torus that passes through the given phase-space point. This procedure will be more costly than that used in Section 4.1 by a factor of a few because several tori will have to be evaluated for each sampling point \mathbf{u}' , but the procedure will yield the same precision as was achieved in Section 5.

7 CONCLUSIONS

A fundamental task of Galactic astronomy is determination of the Galaxy’s gravitational potential Φ because a knowledge of Φ is required for any investigation of the dynamics or evolution of the Galaxy. In Paper I we showed that models constructed from orbital tori can be used to constrain the Galaxy’s DF to good precision from a catalogue that contains only $\sim 10\,000$ stars. In Section 4 we extended this approach to the determination of Φ . Although the extension is straightforward, we found that it is in practice a notable failure. We traced the problem to Poisson noise arising from the use of a finite number of tori in the analysis. The noise level increases with the completeness and precision of the data because the number of tori that contribute significantly to the likelihood of a given star decreases with the volume of the star’s error ellipsoid. We showed that to beat this noise

down to an acceptable level by brute \sqrt{N} growth one would have to use a number of tori that exceeded the number of stars we were considering (10 000) by at least four orders of magnitude.

Torus modelling is an extension of Schwarzschild modelling, so any problem inherent in torus modelling will be shared by Schwarzschild modelling – for a detailed comparison of the two techniques see Binney & McMillan (2011). Made-to-measure modelling (M2M: Bissantz et al. 2004; Dehnen 2009; Morganti & Gerhard 2012) is a modification of Schwarzschild modelling in which one does not hold entire orbits in memory, and as such will suffer badly from discreteness noise when used with data that are complete and/or precise. Straight N-body modelling has the same problems with discreteness noise that M2M modelling has, and in addition extreme difficulty in adapting the model to fit the data. Thus the discreteness noise we exhibited in Section 4 is a major issue for all galaxy-modelling strategies that are based on orbits.

In Section 5 we showed that discreteness noise can be mastered if we evaluate actions as functions of (\mathbf{x}, \mathbf{v}) rather than the other way round. This is very similar to the approach used by Ting et al. (2012), though our consideration of more realistic selection effects and non-negligible observational uncertainty forces us to deal more carefully with the discreteness noise in this case as well. This approach works because we only require *ratios* of likelihoods, and the discreteness noise will cancel from these ratios if we evaluate both likelihoods using the same phase-space points (\mathbf{x}, \mathbf{v}) for the Monte-Carlo sums with which we approximate integrals. With this approach we were able to achieve the numerical precision to determine the scaleheight of the potential almost as accurately as the data allows, which is to within 20 to 30 pc for the catalogues of 10 000 stars that we consider.

Here we failed in our attempt to constrain the potential with tori but this failure does not indicate that tori should be abandoned. Right now they are invaluable for generating models, and in the future they will play a key role in forthcoming models of the chemodynamical evolution of the Galaxy – recall that our analysis here does not recognise chemically distinct populations, which are in reality central to studies of the structure and history of our Galaxy. Moreover, it is possible to upgrade our torus machine so it can evaluate $\mathbf{J}(\mathbf{x}, \mathbf{v})$. The brute-force way to do this is simply to construct tori iteratively until one has constructed the one that passes through a given phase-space point (\mathbf{x}, \mathbf{v}) (McMillan & Binney 2008), but a computationally much faster technique may be possible: currently we exploit each torus in isolation and given observables \mathbf{u} for a star, we find the phases $\boldsymbol{\theta}$ at which a star on a given torus \mathbf{J} passes close to \mathbf{u} . By interpolating between tori it should be possible to find the $(\boldsymbol{\theta}, \mathbf{J})$ combination that brings a star to \mathbf{u} .

ACKNOWLEDGEMENTS

We thank John Magorrian and the other members of the Oxford dynamics group for valuable comments. This work is supported by grants ST/G002479/1 and ST/J00149X/1 from the Science and Technology Facilities Council.

REFERENCES

- Binney J., 2010, MNRAS, 401, 2318
 —, 2011, Pramana, 77, 39
 —, 2012a, MNRAS, 426, 1324
 —, 2012b, MNRAS, 426, 1328
 Binney J., McMillan P., 2011, MNRAS, 413, 1889
 Binney J., Merrifield M., 1998, Galactic astronomy. Princeton University Press
 Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
 Bissantz N., Debattista V. P., Gerhard O., 2004, ApJL, 601, L155
 Bovy J., Tremaine S., 2012, ApJ, 756, 89
 Chanamé J., Kleyna J., van der Marel R., 2008, ApJ, 682, 841
 Creze M., Chereul E., Bienayme O., Pichon C., 1998, A&A, 329, 920
 de Zeeuw T., 1985, MNRAS, 216, 273
 Dehnen W., 2009, MNRAS, 395, 1079
 Dehnen W., Binney J., 1998, MNRAS, 294, 429
 Eisenstein D. J., Weinberg D. H., Agol E., Aihara H., Allende Prieto C., Anderson S. F., Arns J. A., Aubourg É., Bailey S., Balbinot E., et al., 2011, AJ, 142, 72
 Garbari S., Liu C., Read J. I., Lake G., 2012, MNRAS, 425, 1445
 Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, ApJ, 692, 1075
 Gilmore G., Randich S., Asplund M., Binney J., Bonifacio P., Drew J., Feltzing S., Ferguson A., Jeffries R., Micela G., Negueruela I., Prusti T., Rix H.-W., Vallenari A., Alfaro E., Allende-Prieto C., Babusiaux C., Bensby T., Blomme R., Bragaglia A., Flaccomio E., François P., Irwin M., Koposov S., Korn A., Lanzafame A., Pancino E., Paunzen E., Recio-Blanco A., Sacco G., Smiljanic R., Van Eck S., Walton N., 2012, The Messenger, 147, 25
 Holmberg J., Flynn C., 2004, MNRAS, 352, 440
 Holmberg J., Nordström B., Andersen J., 2009, A&A, 501, 941
 Klypin A., Zhao H., Somerville R. S., 2002, ApJ, 573, 597
 Kuijken K., Gilmore G., 1989, MNRAS, 239, 605
 —, 1991, ApJL, 367, L9
 Magorrian J., 2006, MNRAS, 373, 425
 —, 2013, MNRAS submitted, arXiv:1303.6099
 Malhotra S., 1994a, ApJ, 433, 687
 —, 1994b, ApJ, 433, 687
 —, 1995, ApJ, 448, 138
 McMillan P. J., 2011, MNRAS, 414, 2446
 —, 2013, MNRAS, 430, 3276
 McMillan P. J., Binney J. J., 2008, MNRAS, 390, 429
 —, 2010, MNRAS, 402, 934
 —, 2012, MNRAS, 419, 2251
 Morganti L., Gerhard O., 2012, MNRAS, 422, 1571
 Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
 Perryman M. A. C., de Boer K. S., Gilmore G., Høg E., Lattanzi M. G., Lindegren L., Luri X., Mignard F., Pace O., de Zeeuw P. T., 2001, A&A, 369, 339
 Reid M. J., Brunthaler A., 2004, ApJ, 616, 872
 Sanders J., 2012a, MNRAS, 426, 128
 —, 2012b, MNRAS, 425, 2228
 Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403,

1829

- Schwarzschild M., 1979, ApJ, 232, 236
 Steinmetz M., Zwitter T., Siebert A., et al., 2006, AJ, 132, 1645
 Ting Y.-S., Rix H.-W., Bovy J., van de Ven G., 2012, MNRAS submitted, arXiv:1212.0006
 van den Bosch R. C. E., van de Ven G., Verolme E. K., Cappellari M., de Zeeuw P. T., 2008, MNRAS, 385, 647

APPENDIX

In Table 4 we give the parameters of the gravitational potential models with varying R_d used for the tests shown in Figures 7, 9 & 10. In each case the the potential is referred to in the text by its thin disc scalelength.

$\Sigma_{d,0,\text{thin}}$ ($M_{\odot} \text{ pc}^{-2}$)	$R_{d,\text{thin}}$ (kpc)	$z_{d,\text{thin}}$ (kpc)	$\Sigma_{d,0,\text{thick}}$ ($M_{\odot} \text{ pc}^{-2}$)	$R_{d,\text{thick}}$ (kpc)	$z_{d,\text{thick}}$ (kpc)	$\rho_{b,0}$ ($M_{\odot} \text{ pc}^{-3}$)	$\rho_{h,0}$ ($M_{\odot} \text{ pc}^{-3}$)	r_h (kpc)
740.3	2.25	0.30	161.7	2.62	0.90	86.1	0.0271	12.5
704.1	2.50	0.30	163.4	2.92	0.90	86.5	0.0214	13.8
851.3	2.75	0.30	199.8	3.21	0.90	92.4	0.0127	17.0
753.0	3.00	0.30	182.0	3.50	0.90	94.1	0.0125	17.0
673.8	3.25	0.30	165.0	3.79	0.90	98.7	0.0232	11.7
505.8	3.50	0.30	128.5	4.08	0.90	99.7	0.0322	10.4
410.5	3.75	0.30	103.7	4.38	0.90	99.4	0.0613	7.5

Table 4. Parameters of the potentials with varying R_d used in this paper (including the Mc11 potential, with $R_{d,\text{thin}} = 3.0 \text{ kpc}$). In each case the parameters were fit in the same way as Mc11, as described in McMillan (2011) with the assumed Solar radius $R_0 = 8.5 \text{ kpc}$ and with the ratio of the thick and thin disc scalelengths fixed at 3.5/3.